On Matrix Majorants and Minorants, with Applications to Differential Equations

Germund Dahlquist
Department of Numerical Analysis and Computing Science
Royal Institute of Technology
S-100 44 Stockholm 70, Sweden

Dedicated to Professor A. M. Ostrowski on his ninetieth birthday.

Submitted by Walter Gautschi

ABSTRACT

Some tools of linear algebra are collected and developed for potential use in the analysis of stiff differential equations. Bounds for the triangular factors of a large matrix are given in terms of the triangular factors of an associated "minorant" matrix of lower order. Minorants are also used to produce estimates of solutions of systems of ordinary differential equations, which may be sharper than those obtained by the use of logarithmic norms.

1. DEFINITIONS. MINORANTS FOR LINEAR SYSTEMS

Consider a partitioned matrix,

$$A = [A_{ij}]_{ij=1}^{m}, \tag{1.1}$$

where A_{ij} is an $n_i \times n_j$ matrix, $n = \Sigma n_i$. Let $\|\cdot\|$ be any vector norm in $\mathbb{R}^\nu$, $\nu = n_i$, $i = 1,2,\ldots,m$. The same notation is used also for the corresponding operator norms for the rectangular submatrices.

We shall associate $m \times m$ matrices with A in two different ways; see also Ostrowski [12], who used these concepts (but not the same terminology) in order to derive certain determinant inequalities.

$\hat{A} = [\hat{a}_{ij}]$ is called an $m \times m$ *majorant* of A iff

$$\hat{a}_{ij} \geqslant \|A_{ij}\|, \qquad i, j = 1,2,\ldots,m. \tag{1.2}$$

52 Vanderbilt Ave., New York, NY 10017 0024-3795/83/$3.00

Evidently $|\alpha|\hat{A}$, $\hat{A}+\hat{B}$, and $\hat{A}\cdot\hat{B}$ are majorants of, respectively, αA, $A+B$, and $A\cdot B$. The majorant relation is expressed by the inequality

$$A \leqslant \hat{A} \qquad (\text{if} \quad m < n). \tag{1.2'}$$

Inequalities between vectors and matrices of the same type are to hold elementwise: $x \geqslant 0$ means that $x_i \geqslant 0$, and $x < 0$ means that $x_i < 0$, for all i.

$\check{A} = [a_{ij}]$ is called an $m \times m$ *minorant* of A iff

$$a_{ii} \leqslant \|A_{ii}^{-1}\|^{-1}, \qquad a_{ij} \leqslant -\|A_{ij}\|. \tag{1.3}$$

Our definition of a minorant differs from that of Robert [13]. The two concepts come closer with the additional M-matrix condition of Proposition 1. This condition is, however, relaxed in the Corollary and not used at all in the applications to differential equations in Section 3.

The purpose of this paper is to collect and develop tools for potential use in the analysis of methods for stiff ODEs. Some types of applications are indicated (for $m = 2$) by Söderlind and Dahlquist [18].

PROPOSITION 1. *Let $\check{A}$ be an $m \times m$ minorant of the $n \times n$ matrix A. Assume that $\check{A}$ is a nonsingular M-matrix, i.e.*

$$a_{ij} \leqslant 0 \quad (i \neq j), \qquad \check{A}^{-1} \geqslant 0.$$

Then $\check{A}$ has a triangular factorization, $\check{A} = \check{L}\check{U}$, such that $\check{L}$ has unit diagonal elements and $\check{U}$ has positive diagonal elements. A has a unique block triangular factorization, $A = LU$, with unit matrices in the block diagonal of L. $\check{L}, \check{U}$ are minorants of L, U. The inverses of $\check{L}$, $\check{U}$, and $\check{A}$ are majorants of, respectively, the inverses of L, U, and A.

REMARK. There are several equivalent forms for the assumption that $\check{A}$ is a nonsingular M-matrix: see Varga [19], Seneta [15, Chapter 2, in particular Exercise 2.4], Berman and Plemmons [1, especially Chapter 6], and Fiedler and Pták [7].

(1) We can write $\check{A} = aI - T$, where $a \in \mathbb{R}$, $T \geqslant 0$. Then $\check{A}$ is an M-matrix if and only if the spectral radius (the Perron root) of T is less than a.

(2) $\check{A}$ is a nonsingular M-matrix if and only if $\Delta_i > 0$, where Δ_i is the principal minor of $\check{A}$ which consists of its first i rows and columns, $i = 0, 1, 2, \ldots, m$. (We set $\Delta_0 = 1$.)

(3) Let D be the diagonal part of $\check{A}$. $\check{A}$ is an M-matrix iff $D > 0$ and the spectral radius of $I - D^{-1}\check{A}$ is less than one.

(4) All eigenvalues of $\check{A}$ have strictly positive real parts.

See also the generalization in our corollary below.

Proof. Put

$$\check{A} = [a_{ij}], \qquad \check{L} = [l_{ij}], \qquad \check{U} = [u_{ij}],$$

i.e.

$$\sum_{\nu} l_{i\nu} u_{\nu j} = a_{ij} \qquad [\nu \leqslant \min(i, j)].$$

By Remark 2, $\check{L}$ and $\check{U}$ exist, with

$$l_{ii} = 1, \qquad u_{ii} = \frac{\Delta_i}{\Delta_{i-1}} > 0.$$

(This is well known; see e.g. Fiedler and Pták [7]). In analogous notation, we have for A

$$L_{ii} = I, \qquad \sum_{\nu} L_{i\nu} U_{\nu j} = A_{ij} \quad [\nu \leqslant \min(i, j)].$$

We are to show that L and U are defined by these relations and that, for $i \neq j$,

$$\|L_{ij}\| \leqslant -l_{ij}, \qquad \|U_{ij}\| \leqslant -u_{ij}, \qquad \|U_{jj}^{-1}\| \leqslant u_{jj}^{-1}.$$

The proof is by induction, with respect to both subscripts, in the same order as in the actual LU-factorization. (For $m = 1$, the proposition is trivially true.) Note that $l_{i\nu} u_{\nu j} \geqslant 0$, $\nu < \min(i, j)$.

We obtain, for $i = 1, 2, \ldots, j-1$, by the induction hypothesis and the triangle inequality,

$$\|U_{ij}\| = \left\| A_{ij} - \sum_{\nu=1}^{i-1} L_{i\nu} U_{\nu j} \right\|$$

$$\leqslant -a_{ij} + \sum l_{i\nu} u_{\nu j} = -u_{ij}.$$

For $j=1,2,\ldots,i-1$, we have

$$\|L_{ij}\| = \left\| \left(A_{ij} - \sum_{\nu=1}^{j-1} L_{i\nu}U_{\nu j} \right) U_{jj}^{-1} \right\|$$

$$\leqslant \left(-a_{ij} + \sum l_{i\nu}u_{\nu j} \right) u_{jj}^{-1} = -l_{ij},$$

$$\|U_{jj}^{-1}\| = \left\| \left(A_{jj} - \sum_{\nu=1}^{j-1} L_{j\nu}U_{\nu j} \right)^{-1} \right\|$$

$$= \left\| \left(I - A_{jj}^{-1} \sum_{\nu=1}^{j-1} L_{j\nu}U_{\nu j} \right)^{-1} A_{jj}^{-1} \right\|$$

$$\leqslant \left(1 - a_{jj}^{-1} \sum l_{j\nu}u_{\nu j} \right)^{-1} a_{jj}^{-1} = u_{jj}^{-1}.$$

Hence, L, U are uniquely defined, and have $\check{L}, \check{U}$ as minorants.

Now, set

$$L = I_n - L', \qquad \check{L} = I_m - (\check{L})',$$

where $L', (\check{L})'$ are strictly lower triangular. Note that $(\check{L})'$ is a *majorant* of L' and that $(L')^i = 0$ for $i > m$. Now

$$L^{-1} = \sum_{i=0}^{m} (L')^i \leqslant \sum_{i=0}^{m} (\check{L}')^i = (\check{L})^{-1},$$

i.e., the inverse of $\check{L}$ is an $m \times m$ majorant of the inverse of L. Let $D, \check{D}$ be the (block) diagonal parts of the upper (block) triangular matrices $U, \check{U}$. Then we can set

$$D^{-1}U = I_n - V, \qquad (\check{D})^{-1}\check{U} = I_m - W,$$

where W and $(\check{D})^{-1}$ are *majorants* of, respectively, V and D^{-1}, since, for $i \neq j$, $\|V_{ij}\| \leqslant \|U_{ii}^{-1}\| \cdot \|U_{ij}\| \leqslant u_{ii}^{-1}|u_{ij}| = w_{ij}$. Thus

$$U^{-1} = (I_n - V)^{-1}D^{-1} = \sum_{i=1}^{m} V^i D^{-1} \leqslant \sum_{i=1}^{m} W^i (\check{D})^{-1}$$

$$\leqslant (I_m - W)^{-1}(\check{D})^{-1} = (\check{U})^{-1}.$$

Finally,

$$A^{-1} = U^{-1}L^{-1} \leqslant (\check{U})^{-1}(\check{L})^{-1} = (\check{A})^{-1}. \qquad \blacksquare$$

The following generalization is useful for partitioned systems of stiff ODEs.

COROLLARY. *Let Δ_i have the same meaning as in Remark 2. If $\Delta_i > 0$, for $i < m$, but $\Delta_m \leqslant 0$, then, by our induction proof, the factorization $A = LU$ still exists, although $u_{mm} \leqslant 0$. Now $\check{L}, \check{U}$ are minorants of L, U. The existence of the inverses of U and A is, however, no longer guaranteed. The proposition is applicable in full to the leading submatrices of $\check{A}$ and the corresponding submatrices of A.*

The LU-factorization of singular M-matrices has recently been studied; see Funderlic and Plemmons [8], Varga and Cai [20].

Minorants can also be used in the study of iterative processes, e.g. the Gauss-Seidel method. This is a particular case of the theory of iteration in pseudometric spaces, developed by Schröder [14]; see also Collatz [3, Chapter 2]. Such applications are beyond the scope of our paper. Ström [16] applies, however, similar ideas to a problem related to stiff ODE's.

EXAMPLE. Let the $n \times n$ matrix A be block tridiagonal, the typical block row being

$$\ldots, -I, D, -I, \ldots,$$

where D is the tridiagonal $p \times p$ matrix, $p = n/m$, with the typical row

$$\ldots, -1, 4, -1, \ldots.$$

It is well known that

$$\|D^{-1}\|_2^{-1} = 2 + \varepsilon_p, \qquad \varepsilon_p \downarrow 0 \qquad \text{when } p \to \infty.$$

An $m \times m$ minorant $\check{A}$ is then easily found, namely the tridiagonal matrix with the typical row

$$\ldots, -1, 2 + \varepsilon_p, -1, \ldots.$$

The proposition shows how it is possible to obtain bounds for the pivots etc. in the application of block Gauss elimination to A, by performing the computations on a smaller matrix $\check{A}$. The technique can be used in more complicated elliptic problems and may lead to suggestions concerning the choice of pivots etc.

2. SOME BOUNDS FOR NORMS AND EIGENVALUES

Set

$$\begin{aligned} &x = \left(x_1^T, x_2^T, \ldots, x_m^T\right)^T, \qquad x \in \mathbb{R}^n, \quad x_i \in \mathbb{R}^v, \ v = n_i, \\ &\hat{x} = \left(\|x_1\|, \|x_2\|, \ldots, \|x_m\|\right)^T, \qquad \hat{x} \in \mathbb{R}^m. \end{aligned} \tag{2.1}$$

$\hat{x} \in \mathbb{R}^m$ defines a pseudometric in $\mathbb{R}^n$; see e.g. Collatz [3, Chapter 2]. Any vector $y \in \mathbb{R}^m$ such that $y \geqslant \hat{x}$ is called a majorant (m-majorant) of $x \in \mathbb{R}^n$. We write $y \geqslant x$. Let $|\cdot|$ be any absolute and monotone norm in $\mathbb{R}^m$.

For $x \in \mathbb{R}^n$, set

$$\|x\| = |\hat{x}|. \tag{2.1'}$$

Then, if $\check{A}$ is an M-matrix,

$$A^{-1} \leqslant \check{A}^{-1} \quad \Rightarrow \quad \|A^{-1}\| \leqslant |\check{A}^{-1}|. \tag{2.2}$$

Evidently we also have

$$\|A\| \leqslant |\hat{A}|. \tag{2.2'}$$

In particular, let

$$|\xi| = |\xi|_w = \max \frac{|\xi_i|}{w_i}, \qquad \|x\|_w = |\hat{x}|_w,$$

where $w = (w_1, w_2, \ldots, w_m)^T$ is a vector of strictly positive weight factors. Then

$$\|A\|_w \leqslant |\hat{A}|_w = \max_i \frac{(\hat{A}w)_i}{w_i}. \tag{2.3}$$

The infimum, over the set of weight vectors w, of the maximum on the right-hand side is equal to $\max|\lambda(\hat{A})|$, which is a positive eigenvalue of $\hat{A}$, the Perron root. The infimum is assumed for the corresponding eigenvector (the Perron vector), which is nonnegative. It is strictly positive if A is irreducible.[1] Since $\|A\|_w \geqslant \max|\lambda(A)|$,

$$\max|\lambda(A)| \leqslant \max|\lambda(\hat{A})| \leqslant \max_i \frac{(\hat{A}w)_i}{w_i},$$

(Here w is any positive vector.) Similarly, since we may write $\check{A} = aI - T$, we obtain, if $|T|_w < a$,

$$\|A^{-1}\|_w^{-1} \geqslant |\check{A}^{-1}|_w^{-1} = a - |T|_w = \min_i \frac{(\check{A}w)_i}{w_i}. \tag{2.4}$$

The supremum of this minimum is equal to the smallest eigenvalue of $\check{A}$, which is positive if $\check{A}$ is an M-matrix (see the remark to the proposition). The supremum is assumed when w is the corresponding eigenvector, which is nonnegative. We therefore obtain, *if $\check{A}$ is an M-matrix*,

$$\min|\lambda(A)| \geqslant \min|\lambda(\check{A})| \geqslant \min_i \frac{(\check{A}w)_i}{w_i}.$$

A lower bound for the real part of the eigenvalues can be obtained by a special type of minorants; see the next section. See also Feingold and Varga [6] concerning Gerschgorin-type theorems for block matrices.

3. MINORANTS, LOGARITHMIC NORMS, AND ODES

We note that if A is replaced by $A + \alpha I$, the conditions for majorants and minorants are changed in the main diagonal only. We easily find, by the triangle inequality, that for $\alpha > 0$

$$\alpha I_m + \text{any majorant of } A - \alpha I$$

is a majorant of A.

[1] Irreducibility is not necessary for strict positivity; see end of this paper.

If we apply the fact that $\|B^{-1}\|^{-1} = \inf\|Bx\|$, $\|x\| = 1$, for $B = A_{ii}$, we similarly find that for $\alpha > 0$

$$-\alpha I_m + \text{any minorant of } A + \alpha I$$

is a minorant of A. Set $B = A_{ii}$, $\alpha = 1/\varepsilon$, and let $\alpha \to \infty$. Then

$$\|(B+\alpha I)^{-1}\|^{-1} - \alpha = \frac{1-\alpha\|(B+\alpha I)^{-1}\|}{\|(B+\alpha I)^{-1}\|} = \frac{1-\|(I+\varepsilon B)^{-1}\|}{\varepsilon\|(I+\varepsilon B)^{-1}\|}$$

$$= \frac{\|I-\varepsilon B\|-1+O(\varepsilon^2)}{-\varepsilon} \to -\mu(-B),$$

where $\mu(\cdot)$ is the *logarithmic matrix norm* (sometimes called the "measure" of a matrix), introduced and studied by Lozinskii [10] and Dahlquist [4]. (See also Ström [17], Nevanlinna [11].)

It follows that *any matrix* $\check{A} = [a_{ij}]$ *that satisfies the relations*

$$a_{ij} \leqslant \begin{cases} -\mu(-A_{ii}), & i = j, \\ -\|A_{ij}\|, & i \neq j, \end{cases} \tag{3.1}$$

is a minorant of A. These minorants have some nice special properties. It follows e.g. from the subadditivity and homogeneity properties of the logarithmic norms, i.e. $\mu(-A-B) \leqslant \mu(-A)+\mu(-B)$, $\mu(\alpha A) = \alpha\mu(A)$ for $\alpha > 0$, that *if* $\check{A}$, $\check{B}$, *are minorants of* A, B, *satisfying* (3.1), *then* $\check{A}+\check{B}$ *is a minorant of* $A+B$ *and* $\alpha\check{A}$ *is a minorant of* αA, *for any* $\alpha > 0$. In particular, $\check{A}+\alpha I$ is a minorant of $A+\alpha I$ for any $\alpha \in \mathbb{R}$ (also if $\alpha < 0$).

Dahlquist [4] associated matrices of this type to systems of nonlinear differential equations. Consider the nonlinear system[2]

$$\frac{dx}{dt} + A(t,x)\cdot x = p(t,x). \tag{3.2}$$

Let $\check{A}(t)$ *satisfy* (3.1) *for all* $x \in D_r = \{x : \|x\| \leqslant r\}$. Then it is well known, see

[2]This is a convenient form in the study of the difference between the solutions of two neighboring initial value problems.

e.g. Dahlquist [4, p. 16], that for $i = 1,2,\ldots,m$,

$$\frac{d\|x_i\|}{dt} \leqslant \mu(-A_{ii}(t,x))\|x_i\| + \sum_{j\neq i}' \|A_{ij}(\cdot)\|\|x_j\| + \|p_i(t,x)\|,$$

i.e., if we define $\hat{x}, \hat{p}(t,x)$ by (2.1) and let $\hat{p}(t) \geqslant \hat{p}(t,x)$ for all $x \in D_r$, it follows that

$$\frac{d\hat{x}}{dt} + \check{A}(t)\cdot\hat{x} \leqslant \hat{p}(t). \tag{3.3}$$

For $m = 1$, we have the usual differential inequality for $\|x\|$, where $a(t) \leqslant -\mu(-A(t,x))$, $\hat{\pi}(t) \geqslant \|p(t,x)\|$, namely

$$\frac{d\|x\|}{dt} + a(t)\|x\| \leqslant \hat{\pi}(t). \tag{3.3'}$$

PROPOSITION 2. *Let $B(t)$ be a matrix with nonnegative off-diagonal elements, which are continuous functions of t, $t \geqslant t_0$. Let $Y(t)$ be the solution of the initial-value problem*

$$\frac{dY}{dt} = B(t)Y, \qquad Y(t_0) = I_m.$$

Then $Y(t)Y(s)^{-1} \geqslant 0$ for $t \geqslant s \geqslant t_0$, and $Y(t)Y(s)^{-1}$ has no row of zeros.

If, in addition, $B(s')$ is irreducible for some s', $s \leqslant s' < t$, then $Y(t)Y(s)^{-1} > 0$.

Proof. We can assume, without loss of generality, that $B(t)$ is nonnegative, because $Y(t) > 0 \Leftrightarrow Y(t)e^{\alpha t} > 0$ for any $\alpha \in R$, and

$$\frac{d(Ye^{\alpha t})}{dt} = (B + \alpha I)Ye^{\alpha t},$$

and we can choose α so that $B + \alpha I \geqslant 0$ for $t \in [t_0, T]$. Without loss of generality, we can also put $s = t_0 = 0$, since for a general s, we have the same problem for $Y(t)Y(s)^{-1}$ with a shifted time variable.

Consider $y(t) = Y(t)\eta$, where η is an arbitrary positive vector. Then

$$\frac{dy(t)}{dt} = B(t)y(t), \qquad y(0) = \eta > 0.$$

We claim that $Y(t)\eta > 0\ \forall t > 0$. If this were not true, there would exist a $t' > 0$ such that $y(t) > 0$ for $t < t'$, and hence $dy/dt \geqslant 0$ for $t < t'$, while at least one component vanished at $t = t'$, $y_i(t') = 0$ (say). This would lead to the following contradiction:

$$0 \leqslant \int_0^{t'} y_i'(t)\,dt = 0 - y_i(0) = -\eta < 0.$$

Hence $Y(t)\eta > 0$ for all $t > 0$, $\eta > 0$. This shows that $Y(t)$ has no row of zeros. By continuity $Y(t)\eta \geqslant 0$ for all $t \geqslant 0$, $\eta \geqslant 0$, i.e., $Y(t) \geqslant 0$ for $t > 0$. This proves the first part of the proposition.

Next consider, for $s \leqslant s' < t$, and for some $\delta > 0$,

$$Y(t)Y(s)^{-1} = Y(t)Y(s'+\delta)^{-1} \cdot Y(s'+\delta)Y(s')^{-1} \cdot Y(s')Y(s)^{-1}.$$

The first and the third factors on the right-hand side are nonnegative and have no rows of zeros. We shall prove that, for $t > s'$, $Y(t)Y(s)^{-1} > 0$. It is seen by a moment's reflection that it is sufficient to show that

$$Y(s'+\delta)Y(s')^{-1} > 0,$$

for all sufficiently small positive δ.

Again we can, without loss of generality, put $s' = t(0) = 0$, $Y(0) = I$. Therefore, let $B(0)$ be irreducible. By continuity, there exists a constant irreducible matrix C with zeros in the same positions as $B(0)$, such that $B(t) \geqslant C$ on $[0, \delta]$, for any sufficiently small δ. Put

$$Q(t) = \frac{dY}{dt} - CY(t), \qquad Y(0) = I. \tag{3.5}$$

Note that $Q(t) = [B(t) - C]Y(t) \geqslant 0$. It is well known that $e^{C\tau} > 0$ for all $\tau > 0$; see [15, Theorem 2.6]. The equations in (3.5) are, by Duhamel's formula, equivalent to the equation

$$Y(t) = e^{Ct} + \int_0^t e^{C(t-\tau)}Q(\tau)\,d\tau \geqslant e^{Ct} > 0$$

for $0 < t \leqslant \delta$. Hence $Y(\delta) > 0$, i.e., in the original time variable,

$$Y(s'+\delta)Y(s')^{-1} > 0,$$

and the proof is complete. ∎

COROLLARY 1. *Let* $y(t), z(t) \in \mathbb{R}^m$ *satisfy relations of the form*

$$\frac{dy}{dt} \geqslant B(t)y + q(t),$$

$$\frac{dz}{dt} \leqslant B(t)z + q(t).$$

If $z(0) \leqslant y(0)$, *then* $z(t) \leqslant y(t)$ *for* $t \geqslant 0$.
If $B(s)$ *is irreducible for some* $s \geqslant 0$, *then* $z(t) < y(t)$ *for* $t > s$.

Proof. Put $y(t) - z(t) = u(t)$. Then we can write

$$\frac{du}{dt} = B(t)u + r(t), \qquad r(t) \geqslant 0, \quad u(0) \geqslant 0.$$

Hence, by Duhamel's formula,

$$u(t) = Y(t)u(0) + \int_0^t Y(t)Y(s)^{-1}r(s)\,ds \geqslant Y(t)u(0),$$

and the result follows from Proposition 2. ∎

COROLLARY 2. *For arbitrary* $z(0)$

$$\frac{dz}{dt} \leqslant B(t)z \quad \Rightarrow \quad z(t) \leqslant Y(t)z(0),$$

$$\frac{dz}{dt} \geqslant B(t)z \quad \Rightarrow \quad z(t) \geqslant Y(t)z(0).$$

Proof. These are special cases of Corollary 1. ∎

COROLLARY 3. *If* $dy/dt + \check{A}(t)y \geqslant \hat{p}(t)$, $y(0) \geqslant \hat{x}(0)$, *where* $\check{A}, \hat{p}$ *have the same meaning as in* (3.3), *then* $y(t) \in \mathbb{R}^m$ *is a majorant of* $x(t) \in \mathbb{R}^n$, *where* $x(t)$ *is a solution of* (3.2). *This is valid as long as* $|y(t)| \leqslant r$ [*i.e. as long as the validity of* (3.1) *is guaranteed*].

Proof. Set $B(t) = -\check{A}(t)$, $q(t) = \hat{p}(t)$, $z = x$ in Corollary 1. ∎

COROLLARY 4. *Let $\check{A}$ be a minorant of A, satisfying* (3.1). *Then*

$$\min \operatorname{Re} \lambda(A) \geqslant \min \operatorname{Re} \lambda(\check{A}) \geqslant -\mu(-\check{A}).$$

Proof. Apply Corollary 3 with $A(t, x) = A = \text{const}$, $p(t, x) = 0$, $x(0) = I$, $\hat{x}(0) = I$. Then

$$\|e^{-At}\| = \|x(t)\| = |\hat{x}(t)| \leqslant |e^{-\check{A}t}|.$$

The first assertion now follows from the relation

$$\max \operatorname{Re} \lambda(-A) = \lim_{t \to \infty} t^{-1} \log \|e^{-At}\|,$$

and the analogous relation for $\check{A}$. The second inequality is a well-known property of the logarithmic norm. ∎

LEMMA. *Let $\hat{x}(t)$ satisfy* (3.3), *and let $w(t) > 0$ be a weight vector that satisfies an inequality of the form*

$$\dot{w} + \check{A}w \geqslant \gamma w, \tag{3.4}$$

where γ may depend on t. Then

$$x(t) \leqslant \hat{x}(t) \leqslant \Phi(t)w + Y(t)\hat{x}(0) \leqslant \Psi(t)w,$$

where

$$\dot{\Phi} + \gamma\Phi = |\hat{p}|_w, \qquad \Phi(0) = 0,$$

$$\dot{\Psi} + \gamma\Psi = |\hat{p}|_w, \qquad \Psi(0) = |\hat{x}|_{w(0)},$$

$$\dot{Y} + \check{A}Y = 0, \qquad Y(0) = I_m.$$

REMARK. See also the improvement in Proposition 3.

Proof. Note that $\Phi(t) \geqslant 0$, $\Psi(t) \geqslant 0$ for $t \geqslant 0$. Set

$$y(t) = \Phi(t)w + Y(t)\hat{x}(0).$$

Then

$$\dot{y} + \check{A}y = \dot{\Phi}w + \Phi\dot{w} + \Phi\check{A}w \geqslant (\dot{\Phi} + \gamma\Phi)w$$

$$= |\hat{p}|_w w \geqslant \hat{p}.$$

Since $y(0) = \hat{x}(0)$, then the first result follows from Corollary 3. The second result similarly follows by setting $y(t) = \Psi(t)w$. ■

With given w, the bounds in the lemma become better, the larger γ is chosen. If w is constant, the condition on w, γ is $\check{A}w \geqslant \gamma w$, i.e.

$$\gamma \leqslant \min \frac{(\check{A}w)_i}{w_i} \tag{3.5}$$

This bound is not better than the result obtained by the use of the logarithmic norm $\mu_w(\cdot)$ subordinate to the $\|\cdot\|_w$ norm. For it is well known that for $C = [c_{ij}]_{i,j=1}^m$,

$$\mu_w(C) = \max_i \left(c_{ii} + \sum_{j \neq i}' \frac{|c_{ij}|w_j}{w_i} \right).$$

Hence,

$$-\mu_w(-\check{A}) = \min \frac{(\check{A}w)_i}{w_i}, \tag{3.6}$$

i.e. the same as the bound for γ. Analogously to (2.2′), one can show that

$$-\mu_w(-A) \geqslant -\mu_w(-\check{A}). \tag{3.6′}$$

It was pointed out in Section 2 that

$$\sup_{w>0} \min \frac{(\check{A}w)_i}{w_i}$$

is obtained when w is the eigenvector corresponding to the Perron root of the non-negative matrix $T = aI - \check{A}$. This vector will be called the Perron vector and denoted w_{Per}. We shall return to the choice of w at the end of the paper.

In the derivation of the lemma, the use of the inequality

$$\hat{p} \leqslant |\hat{p}|_w w$$

causes a loss of sharpness. A w-norm which gives a large γ may not give an adequate description of the vector $\hat{p}$. Note e.g. that if $\check{A}$ is a constant M-matrix, then $\hat{x}(t) \leqslant \check{A}^{-1}\hat{p}$ when $t \to \infty$. $\check{A}$ is called a *graded matrix* if a positive diagonal matrix E with elements of very different orders of magnitude can be found, such that $B = E\check{A}$ has elements of "normal" size. (Such matrices occur in many applications of ODEs.) Then $\check{A}^{-1}\hat{p} = B^{-1}E\hat{p}$, showing that $E\hat{p}$ tells the relative importance of the elements of p.

EXAMPLE.

$$\check{A} = \begin{bmatrix} 1000 & -999 \\ 0 & 1 \end{bmatrix}, \qquad \hat{p}\begin{bmatrix} 1+c \\ 1 \end{bmatrix}, \qquad w = w_{\text{Per}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

(w_{Per} = Perron vector of $aI - \check{A}$ for $a \geqslant 1000$). Set $c^+ = \max(c, 0)$. Then

$$\check{A}^{-1} = \begin{bmatrix} 0.001 & 0.999 \\ 0 & 1 \end{bmatrix}, \qquad \check{A}^{-1}\hat{p} = \begin{bmatrix} 1+0.001c \\ 1 \end{bmatrix},$$

$$\check{A}^{-1}|\hat{p}|_w w = (1+c^+)w, \qquad |\check{A}^{-1}\hat{p}|_w w = (1+0.001c^+)w.$$

There is a considerable loss of sharpness when $\hat{p}$ is replaced by $|\hat{p}|_w w$ if $c \gg 1$. ■

We shall suggest a treatment (for the variable-coefficient case) which appears to be better than the logarithmic-norm approach. Consider (3.3),

$$\frac{d\hat{x}}{dt} + \check{A}\hat{x} \leqslant \hat{p},$$

where $\hat{p}$ is a majorant of $p(t, x)$. Since $\check{A}$ may have some small eigenvalues, we shall compare $\hat{x}$ with $q = (\check{A} + \alpha I)^{-1}\hat{p}$, for some suitably chosen α, instead of $\check{A}^{-1}\hat{p}$. Set, therefore, $z = \hat{x} - q$. Then

$$\dot{z} + \check{A}z \leqslant \hat{p} - \dot{q} - \check{A}q = \alpha q - \dot{q} \leqslant (\alpha + \beta)|q|_w w$$

if β is chosen so that

$$\dot{q} \geqslant -\beta|q|_w w. \tag{3.7}$$

By the lemma, we then obtain the following result.

PROPOSITION 3. *Let $\hat{x}(t)$ be a majorant of $x(t)$ that satisfies* (3.3). *Assume that w and γ satisfy* (3.4), *and set $q = (\check{A} + \alpha I)^{-1}\hat{p}$, where α is to be chosen so that $(\check{A} + \alpha I)^{-1} \geqslant 0$. Choose β to satisfy* (3.7).

Then

$$\hat{x}(t) \leqslant q(t) + \Psi(t)w$$

where $\Psi(t)$ is any function satisfying the inequalities

$$\dot{\Psi} + \gamma\Psi \geqslant (\alpha + \beta)|q|_w, \qquad \Psi(0) \geqslant |\hat{x}(0) - q(0)|_{w(0)}.$$

We shall not discuss here the choice of $\hat{p}$, α, and β. Let us only mention that (3.7) has two advantages to the more obvious alternative to use a logarithmic norm after an analogous transformation of (3.2), instead of (3.3), replacing q, β by $\tilde{q}, \tilde{\beta}$ defined as follows:

$$\tilde{q} = (A + \alpha I)^{-1} p(t, x), \qquad \|\dot{\tilde{q}}\| \leqslant \tilde{\beta}\|\tilde{q}\|, \tag{3.8}$$

First, $\hat{p}(t)$ and $\check{A}(t)$ may be chosen to be smoother functions of t than, respectively, $p(t, x)$ and $A(t, x)$; hence $\|\dot{q}\|$ may be smaller than $\|\dot{\tilde{q}}\|$. Second, (3.7) is a one-sided bound, while the inequality in (3.8) is not.

REMARK. In the multiplication by $(\check{A} + \alpha I)^{-1}$ the components of $\hat{p}$ in the direction of w_{Per} are strongly amplified; hence less information is lost in the use of the inequality $q \leqslant |q|_w w$ than in $\hat{p} \leqslant |\hat{p}|_w w$, when w is close to w_{Per}. (See the example above.)

Finally, we return to the choice of w. It was pointed out earlier than $w = w_{\text{Per}}$ would be a good choice from the point of view of making γ as large as possible. There are, however, two things to be said about the Perron vector. First, if $\check{A}$ depends on t, usually w_{Per} does so too, and then $\dot{w}$ has to be taken into account in (3.4). This corresponds to the correction which has to be made to the logarithmic-norm estimate, when a time-dependent norm is used. (This is dealt with in another report in preparation.)

The second comment is that the Perron vector may not be strictly positive, when $T = aI - \check{A}$ is reducible, so that the vector norm $|\cdot|_w$ is not defined.

If T is close to a reducible matrix, then it may happen that $(\max w_i)/(\min w_i) \gg 1$ and the w-norm is ill conditioned relative to the

max-norm. This is practically relevant, e.g. for (almost) triangular matrices. Let us look at the case $m = 2$. A reducible 2×2 matrix can be brought (by similarity permutation) to the form

$$T = \begin{bmatrix} b & c \\ 0 & d \end{bmatrix}, \qquad \check{A} = \begin{bmatrix} a-b & -c \\ 0 & a-d \end{bmatrix}.$$

(Note that $T \geqslant 0$.) If $b \geqslant d$, then $w_{\text{Per}} = (1,0)^T$, and if $c \neq 0$, no positive vector satisfies $Tw \leqslant \mu w$ for $\mu \leqslant b$. *If* $d > b$, however, $w_{\text{Per}} = (c, d-b)^T$. *In this case,* $w_{\text{Per}} > 0$ *in spite of the reducibility of* T. Moreover, every vector $w = (u, v)$ such that $u \geqslant cv/(d-b) > 0$ satisfies the inequality $Tw \leqslant dw$. The Perron vector defines a norm, which is ill conditioned relative to the max norm if $c \ll d - b$, but the max norm itself can also be used in this case. The Perron vector defines an ill-conditioned norm (relative to the max norm) also if $c \gg d - b$, and in this case we have to use a γ which is less than $a - d$. Kreiss [9] also points out difficulties in this case.

In a typical stiff case, we may have (say),

$$\check{A} = \begin{bmatrix} 1000 & -c \\ 0 & 1 \end{bmatrix}, \qquad T = \begin{bmatrix} 0 & c \\ 0 & 999 \end{bmatrix}, \qquad w_{\text{Per}} = \begin{bmatrix} 1 \\ 999/c \end{bmatrix}$$

(with $a = 1000$). The Perron root is $\rho_T = 999$, and the corresponding eigenvalue of $\check{A}$ is 1. Only if $c \gg 1000$ do we have to choose γ less than 1. If $c \ll 1000$, we can take $\gamma = 1$, but it seems to be advisable to use the max norm, i.e. $w^T = (1,1)$ instead of w_{Per}.

This discussion can be extended to the case where b and d are square matrices, if their Perron roots satisfy $\rho_d > \rho_b$. Then $Tw \leqslant \rho_d w$ for $w^T = (u^T, v^T)$, if we can choose $v > 0$ such that

$$dv \leqslant \rho_d v$$

and then choose u such that

$$u > 0, \quad bu \leqslant \omega u, \quad u \geqslant \frac{cv}{\rho_d - \omega} \qquad (\rho_b \leqslant \omega < \rho_d).$$

It follows that $\check{A}w \geqslant (a - \rho_d)w$. Note that this is applicable when $\check{A}$ is an upper triangular matrix with nonpositive off-diagonal elements, if the diagonal element of the bottom row of T is strictly larger than the other diagonal elements.

Similarity transformations of A to (almost) block upper triangular form are of interest in the analysis of stiff problems with graded Jacobians, where the large elements are on the top. Dahlquist [5] shows that the block upper triangular matrix will then be close to the matrix U in the factorization treated in Section 1 of the present paper. This is a partial explanation of our interest in the bounds derived in Proposition 1 and its corollary.

This work was carried out during the author's visit at Stanford University, spring 1982, partly sponsored by the NSF Grant No. MCS-7811985. The author is grateful to Gene Golub, Gérard Meurant, and Axel Ruhe for enlightening discussions, and to Gene Golub also for providing excellent working conditions. Thanks are also due to Mr. Mao Zu-fan of the Chinese Academy Computing Centre, Beijing, for his comments after a very careful reading of the manuscript, and to the referees for several valuable references.

REFERENCES

1 A. Berman and R. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic, 1979.

2 A. Bode, Matrizielle untere Schranken linearer Abbildungen und M-Matrizen, *Numer. Math.* 11:405–412 (1968).

3 L. Collatz, *Funktionalanalyse und Numerische Mathematik*, Springer, Berlin, Göttingen, Heidelberg, 1964.

4 G. Dahlquist, *Stability and Error Bounds in the Numerical Integration of Ordinary Differential Equations*, Almqvist & Wiksells Boktryckeri AB, Uppsala, 1958; reprinted as Trans. Royal Inst. Technology, No. 130, Stockholm, 1959.

5 G. Dahlquist, On transformations of graded matrices, with applications to stiff ODEs, TRITA-NA report, Royal Institute of Technology, Stockholm, to appear.

6 D. Feingold and R. Varga, Block diagonally dominant matrices and generalizations of the Gerschgorin theorem, *Pacific J. Math.* 12:1241–1250 (1963).

7 M. Fiedler and V. Pták, On matrices with non-positive off-diagonal elements and positive principal minors, *Czechoslovak Math. J.* 12(87):382–400 (1962).

8 R. F. Funderlic and R. J. Plemmons, LU decomposition of M-matrices by elimination without pivoting, *Linear Algebra Appl.* 41:99–110 (1981).

9 H. O. Kreiss, Difference methods for stiff ordinary differential equations, *SIAM J. Numer. Anal.* 15:21–57 (1958).

10 S. M. Lozinskii, Error estimate for numerical integration of ordinary differential equations, Part I (in Russian), *Izv. Vyssš. Učebn. Zaved. Matematika* 6:52–90 (1958).

11 O. Nevanlinna, On the logarithmic norms of a matrix, Report HTKK-MAT-A94, Helsinki Univ. of Technology, 1976.

12 A. M. Ostrowski, On some metrical properties of operator matrices and matrices partitioned into blocks, *J. Math. Anal. Appl.* 2:161–209 (1961).

13 F. Robert, Recherche d'une M-matrice parmi les minorantes d'un opérateur linéaire, *Numer. Math.* 9:189–199 (1966).
14 J. Schröder, Nichtlineare Majoranten beim Verfahren der Schrittweisen Näherung, *Arch. Math. (Basel)* 7:471–484 (1956).
15 E. Seneta, *Non-negative Matrices*, Wiley, New York, 1973.
16 T. Ström, On the practical application of majorants for nonlinear matrix iterations, *J. Math. Anal. Appl.* 41:137–137 (1973).
17 T. Ström, On logarithmic norms, *SIAM J. Numer. Anal.* 12:741–753 (1975).
18 G. Söderlind and G. Dahlquist, Error propagation in stiff differential systems of singular perturbation type, Report TRITA-NA-8108, Royal Inst. of Technology, Stockholm, 1981.
19 R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N.J., 1962.
20 R. S. Varga and D.-Y. Cai, On the LU factorization of M-matrices, *Numer. Math.* 38:179–192 (1981).

Received 23 July 1982; revised 10 January 1983